

Name: \_\_\_\_\_

1. What is a residual? What does a positive residual tell us? What does a negative residual tell us?

Residual is the difference between the observed  $y$  & the predicted  $y$ .

Positive Residual - The Linear model underestimates

Negative Residual - The Linear model overestimates

2. The Least-Squares Regression Line is the line for which the sum of the squared residuals is the smallest.

3. The LSRL for data that has been standardized (converted to z-scores) is  $\hat{z}_y = r z_x$ .

4. Given the LSRL:  $\hat{y} = 27.6 + .56x$

a. What is the slope of the line? .56

b. What is y-intercept? 27.6

c. What does  $\hat{y}$  represent? ← predicted final exam grade

d. Suppose that you received an 80% on your midterm. What would you expect to get on your final? 72.4 ( $27.6 + .56 \cdot 80$ )

e. What if you received a 75% on your final (you had an 80% on your midterm)?

What is your residual?

2.6

residual =  $y - \hat{y}$

$$= 75 - 72.4 = 2.6$$

5. What three conditions need to be met in order to use a Linear Model?

a. Quantitative Data Condition

b. Straight Enough Condition

c. Outlier Condition

6. A scatterplot of residuals is a good way to check the Straight Enough Condition.

7. If a linear model is appropriate for the data, the residual plot will be BORING. It should not have any unusual features. It should stretch horizontally, with a about the same amount of spread throughout. It should show no bends or outliers.

8. The squared correlation,  $R^2$ , gives the fraction of the data's variation accounted for by the X-variable.

9. An  $R^2$  value of zero means none of the variance is in the data - all of

10. An  $R^2$  value is written as a percent.

it is in the residuals.

→ It means The model does not predict the  $y$ -values - IT IS A HORRIBLE LSRL, whose predictions are completely unreliable.

11. The  $R^2$  value given an indication of how well the model predicts the y-values.

12. If we are looking at a linear model showing the relationship between Midterm Exam Scores (predictor variable) and Final Exam Scores (response variable) and calculated  $R^2$  to be equal to 85%, we could interpret that as meaning: "85% of the variation in final exam scores can be accounted for by variation in midterm exam scores."

13. True or False. The  $R^2$  value determines if the linear model is appropriate for the data.

14. Please look over the Computer Printout for a Regression Analysis on page 188.

$R^2$  tells you -  
✓ "How accurate is the model?"